

Time-inconsistent stochastic optimal control problems in insurance and finance

Lukasz Delong

Institute of Econometrics, Division of Probabilistic Methods

Warsaw School of Economics SGH

Al. Niepodległości 162, 02-554 Warsaw, Poland

lukasz.delong@sgh.waw.pl

Abstract: In this paper we study time-inconsistent stochastic optimal control problems. We discuss the assumption of time-consistency of the optimal solution and its fundamental relation with Bellman equation. We point out consequences of time-inconsistency of the optimal solution and we explain the concept of Nash equilibrium which allows us to handle the time-inconsistency. We describe an extended Hamilton-Jacobi-Bellman equation which can be used to derive an equilibrium strategy in a time-inconsistent stochastic optimal control problem. We give three examples of time-inconsistent dynamic optimization problems which can arise in insurance and finance. We present the solution for exponential utility maximization problem with wealth-dependent risk aversion.

1 Introduction

Stochastic optimal control is now a well-developed field of study, see e.g. Fleming and Rishel (1975), Yong and Zhou (1999), Øksendal and Sulem (2004), Pham (2009). The key assumption in the study of stochastic optimal control problems is the time-consistency of the optimal solution. The time-consistency provides the theoretical foundation for Dynamic Programming Principle and Hamilton-Jacobi-Bellman equation, which are the pillars of the modern stochastic control theory.

Basic optimization problems in insurance and finance are time-consistent. However, one can find a lot of arguments for changing the basic model's assumptions and these changes may lead to a time-inconsistent optimization problem. As the key example, let us consider an investor who maximizes the expected exponential utility from his/her terminal wealth. If the investor's coefficient of risk aversion does not change in time, then the optimization problem is time-consistent and it is known how to derive the optimal solution. However, if we assume that the coefficient of risk aversion changes in time and depends on the investor's available wealth, which seems to be a more reasonable assumption in a dynamic asset allocation problem, then the optimization problem becomes time-inconsistent. The property of time-inconsistency means that the optimal solution and the optimal value function do not satisfy Bellman's Principle of Optimality. Consequently, it is not clear how to define and derive the optimal strategy for a time-inconsistent optimization problem.

Ekeland and Lazrak (2006), Ekeland and Pirvu (2008), Björk and Murgoci (2014) and Björk et al. (2017) developed a theory for solving time-inconsistent optimization problems. They derived an extended version of Hamilton-Jacobi-Bellman equation and introduced a notion of an optimal strategy for time-inconsistent optimization problems which was based on the equilibrium of a game. When solving time-inconsistent optimization problems, we should not search for optimal strategies but we should search for subgame perfect Nash equilibrium strategies.

The goal of this paper is to present general ideas and intuition behind the key notions of time-consistency, time-inconsistency, optimal strategy and equilibrium strategy in stochastic optimal control problems. We do not touch mathematical details, which can be found in the cited literature. This paper should be of interest to researchers in economics, insurance and finance who would like to investigate optimization problems. We discuss the assumption of time-consistency of the optimal solution and its fundamental relation with Bellman equation. We point out consequences of time-inconsistency of the optimal solution and we explain the concept of Nash equilibrium which allows us to handle the time-inconsistency. We describe an extended Hamilton-Jacobi-Bellman equation which can be used to derive an equilibrium strategy in a time-inconsistent stochastic optimal control problem. We give three examples of time-inconsistent dynamic optimization problems which can arise in insurance and finance. We present the solution for exponential utility maximization problem with wealth-dependent risk aversion. As far as we know, this is the first solution derived in an explicit form for exponential utility maximization problem with wealth-dependent risk aversion.

2 Time-consistency and Hamilton-Jacobi-Bellman equation

In this section we study classical stochastic optimal control problems and introduce basic notations. We explain the Bellman's Principle of Optimality and the key property of time-consistency of the optimal solution.

Let us consider a finite time horizon T . As usual, let $(\Omega, \mathbb{F}, \mathbb{P})$ denote a probability space with filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ which is the natural filtration generated by one-dimensional Brownian motion $W := (W(t), 0 \leq t \leq T)$. We investigate a controlled stochastic process $X^\pi : (X^\pi(t), 0 \leq t \leq T)$ which takes the form

$$dX^\pi(t) = \mu(t, X^\pi(t), \pi(t))dt + \sigma(t, X^\pi(t))dW(t), \quad 0 \leq t \leq T, \quad (2.1)$$

where μ and σ denote the drift and the volatility of the process X^π , and π denotes the control strategy. The control strategy π is a stochastic process which satisfies some integrability and measurability assumptions. We assume that the stochastic differential equation (2.1) has a unique strong solution X^π . The process X^π is called a state process.

Example 1. Let us consider a financial market which consists of a risk-free deposit $D = (D(t), 0 \leq t \leq T)$ and a risky stock $S = (S(t), 0 \leq t \leq T)$. The value of the risk-free deposit is constant, $D(t) = 1, 0 \leq t \leq T$, i.e. we assume that the interest rate is zero or we consider discounted quantities in our problem. The price of the risky stock is modelled with the geometric Brownian motion:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t), \quad 0 \leq t \leq T.$$

Let $\pi(t)$ denote the amount of money that the investor/insurer invests at time t in the risky stock S . The wealth process X^π of the investor satisfies the SDE

$$\begin{aligned} dX^\pi(t) &= \pi(t)(\mu dt + \sigma dW(t)) \quad 0 \leq t \leq T, \\ X(0) &= x, \end{aligned}$$

where $x > 0$ denotes the initial wealth. We can see that in this example the control strategy π is interpreted as the amount of money invested in the risky stock, and the controlled stochastic process X^π is interpreted as the wealth of the investor/insurer. The investor/insurer chooses the control strategy π and control its wealth X^π in order to fulfill some objective at the maturity T . We will continue this example in the sequel. \square

We need to introduce the generator of the process (2.1). Let $f(t, x)$ denote a function which is continuously differentiable in t and twice continuously differentiable in x , i.e. $f \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R})$. The generator \mathcal{L} on f is defined by

$$\begin{aligned} \mathcal{L}^\pi f(t, x) &= \lim_{h \rightarrow 0} \frac{\mathbb{E}_{t,x}[f(t+h, X^\pi(t+h))] - f(t, x)}{h} \\ &= f_t(t, x) + \mu(t, x, \pi) f_x(t, x) + \frac{1}{2} \sigma^2(t, x, \pi) f_{xx}(t, x), \end{aligned} \quad (2.2)$$

where $\mathbb{E}_{t,x}[\cdot]$ denotes the conditional expected value $\mathbb{E}[\cdot|X(t) = x]$, and f_t, f_x, f_{xx} denote partial derivatives of f .

Classical stochastic optimal control theory deals with optimization problems of the form

$$\sup_{\pi} \mathbb{E} \left[\int_0^T C(s, X^\pi(s), \pi(s)) ds + G(X^\pi(T)) \right], \quad (2.3)$$

where C and G are interpreted as intermediate and terminal utility, and X^π is given by the SDE (2.1). The control strategy π , which is used to govern the process X^π , can be an arbitrary stochastic process. However, it is common to deal with Markov control strategies and consider control strategies such that the strategy at time t given that $X(t) = x$ is of the form $\pi(t, x)$ where π is a deterministic function. The deterministic function π is called a control policy function or a decision rule. Hence, the solution to the optimization problem (2.3) is a function π which tells what the strategy/decision should be given any possible value of the state process X . In the sequel, π will denote the control strategy (a stochastic process) and the control policy function (a deterministic function).

In order to solve the optimization problem (2.3), we investigate the family of optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T C(s, X^\pi(s), \pi(s)) ds + G(X^\pi(T)) \right]. \quad (2.4)$$

The family (2.4) is indexed with the pair (t, x) which describes the initial time t and the initial state x of the process X^π at time t . Using the Markov property of the control strategy π and the state process X^π , we can introduce the objective function

$$V^\pi(t, x) = \mathbb{E}_{t,x} \left[\int_t^T C(s, X^\pi(s), \pi(s)) ds + G(X^\pi(T)) \right], \quad (2.5)$$

and the optimal value function

$$V(t, x) = \sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T C(s, X^\pi(s), \pi(s)) ds + G(X^\pi(T)) \right]. \quad (2.6)$$

The strategy π^* for which $V^{\pi^*}(t, x) = V(t, x)$ is called the optimal control strategy.

The key property which allows us to solve the optimization problems (2.4) is the Bellman's Principle of Optimality. The Bellman's Principle of Optimality and the Dynamic Programming Principle are the main tools in dynamic optimization and stochastic control theory. The Bellman's Principle of Optimality says that the optimal policy function π^* has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute the optimal policy function with regard to the state resulting from the first decision. In other words, the optimal control policy function π^* which solves the optimization problem (2.4) is independent of the initial pair (t, x) . The policy function π^* which is the optimal solution for the objective (2.5) at (t, x) is still optimal when we solve the same optimization problem (2.4) at some latter point $(s, X^{\pi^*}(s))$. This is a very rational property and this property is called time-consistency of the optimal solution.

As suggested by the principle of optimality, we can separate all future decisions from the current decision. Let π^* denote the optimal control strategy determined by the optimal control policy function π^* . The optimal value function can be written in the following way:

$$V(t, x) = \mathbb{E}_{t,x} \left[\int_t^{t+h} C(s, X^{\pi^*}(s), \pi^*(s)) ds + \int_{t+h}^T C(s, X^{\pi^*}(s), \pi^*(s)) ds + G(X^{\pi^*}(T)) \right]. \quad (2.7)$$

By the Bellman's Principle of Optimality we also have

$$V(t+h, X^{\pi^*}(t+h)) = \mathbb{E} \left[\int_{t+h}^T C(s, X^{\pi^*}(s), \pi^*(s)) ds + G(X^{\pi^*}(T)) | \mathcal{F}_{t+h} \right], \quad (2.8)$$

for any π applied on $[t, t+h]$. Combining the last two equations (2.7)-(2.8) and using the property of conditional expectations, we end up with the

recursive equation for the optimal value function:

$$V(t, x) = \mathbb{E}_{t,x} \left[\int_t^{t+h} C(s, X^{\pi^*}(s), \pi^*(s)) ds + V(t+h, X^{\pi^*}(t+h)) \right]. \quad (2.9)$$

We point out that the recursive equation (2.9) can be derived since the property of time-consistency of the optimal solution holds and the property of conditional expectations can be applied. If one of these properties fails, then the recursive equation of the form (2.9) for the optimal value function cannot be derived.

If we were considering a discrete time model, i.e. a model in which the strategies were chosen at discrete times $0, h, \dots, t, t+h, \dots, T-h$ and kept fixed in between, then the second term in (2.9), i.e. the function $V(t+h, X^{\pi^*}(t+h))$, would depend on the strategies applied at time $t+h, t+2h, \dots, T-h$. It should be intuitively clear that the optimal strategy at time t should be determined by the policy function which solves the following optimization problem:

$$\begin{aligned} \pi^*(t, x) \\ = \arg \max_{\pi} \left\{ \mathbb{E}_{t,x} \left[\int_t^{t+h} C(s, X^{\pi}(s), \pi(s)) ds + V(t+h, X^{\pi}(t+h)) \right] \right\}. \end{aligned} \quad (2.10)$$

Looking at (2.9)-(2.10), we can conclude that the optimization problem (2.3) can be stated in a recursive, step-by-step form. This recursive form describes the relationship between the value function in one period and the value function in the next period. The optimization performed in each period involves maximizing the sum of the period-specific intermediate utility function and the objective function under the optimal strategy at the next period, giving that the strategy is contingent on the value of the state process in the period considered and the optimal decisions made in the next periods. Each period's decision is made by acknowledging that all future decisions will be optimally made. The procedure continues recursively back in time and allows to derive the optimal control strategies. This solution method is called Dynamic Programming Principle. We can see that the Bellman's Principle of Optimality

and the Dynamic Programming Principle transform a dynamic optimization problem into a sequence of one-dimensional optimization problems.

We now move to our continuous-time model. If we divide (2.9) by h , let $h \rightarrow 0$ and use the generator (2.2), we can state the Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} \sup_{\pi} \{ \mathcal{L}^{\pi} V(t, x) + C(t, x, \pi) \} &= 0. \quad (t, x) \in [0, T) \times \mathbb{R}, \\ V(T, x) &= G(x), \quad x \in \mathbb{R}. \end{aligned} \quad (2.11)$$

Let us recall that Hamilton-Jacobi-Bellman equation is Bellman equation in continuous-time models. The optimal control strategy π^* is determined by the control policy functions which solve the optimization problems:

$$\pi^*(t, x) = \arg \max_{\pi} \{ C(t, x, \pi) + \mathcal{L}^{\pi} V(t, x) \}. \quad (2.12)$$

The logic behind the optimization process when we investigate continuous-time models is analogous as in discrete-time models.

Example 2. We continue Example 1. Exponential utility maximization problem is one of the dynamic investment problems often studied in financial and insurance mathematics. The goal is to find the investment strategy which leads to the maximal expected exponential utility from the terminal wealth, i.e. the goal is to solve the optimization problem

$$\sup_{\pi} \mathbb{E} \left[-e^{-\gamma X^{\pi}(T)} \right],$$

where γ denotes the risk aversion coefficient of the investor. Using (2.11), the HJB equation for this optimization problem takes the form

$$\begin{aligned} V_t(t, x) + \sup_{\pi} \left\{ \pi \mu V_x(t, x) + \frac{1}{2} \pi^2 \sigma^2 V_{xx}(t, x) \right\} &= 0, \quad (t, x) \in [0, T) \times \mathbb{R}, \\ V(T, x) &= -e^{-\gamma x}, \quad x \in \mathbb{R}. \end{aligned}$$

We can try to substitute $V(t, x) = -e^{h(t)x+k(t)}$. By doing calculations, we can conclude that the optimal investment strategy is $\pi^*(t) = \frac{\mu}{\gamma\sigma^2}$, see Chapter 5 in Carmona (2009) for details. \square

3 Time-inconsistency and examples in insurance and finance

In this section we explain what time-inconsistency of the optimal solution means. We give three important examples of dynamic optimization problems in insurance and finance which lead to time-inconsistent solutions.

As discussed in the previous section, time-consistency of the optimal solution is a natural property. However, this property does not always hold. We will give specific examples of time-inconsistent optimization problems from insurance and finance in the sequel. First, let us comment on the general idea behind the time-inconsistency. Let us investigate the family of optimization problems of the form

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T C(t, x, s, X^{\pi}(s), \pi(s)) ds + G(t, x, X^{\pi}(T)) \right]. \quad (3.1)$$

Compared to the objective (2.4), the utilities C and G depend now on the initial pair (t, x) which describes the point in time t and the value of the state process at time t . It turns out that the Bellman's Principle of Optimality does not hold for the family of optimization problems (3.1). We can fix and value the utilities in the objective (3.1) at a pre-defined pair (t, x) and use the solution methods described in the previous section. The optimal control policy function π^* which characterizes the optimal solution (the optimal decision rule) for the problem (3.1) for a pre-defined pair (t, x) depends now on this pair (t, x) . The optimal solution is said to be time-inconsistent. Time-inconsistency of the optimal solution means that the policy function π^* which is the optimal solution to the optimization problem for the initial pair (t, x) is no longer optimal at some latter point $(s, X^{\pi}(s))$. In other words, time-inconsistency of the optimal solution means that the restriction of the policy function π^* optimal for the pair (t, x) on a latter time interval $[s, T]$ does not coincide with the policy function π^* optimal for the pair $(s, X^{\pi}(s))$. The optimization process becomes less intuitive. The controller believes that he/she should follow an optimal decision rule which he/she has derived at

time t when the state process is equal to x , but after time t his/her decision rule is no longer optimal and he/she should switch to a different decision rule. Note that when we solve the optimization problem (3.1) by applying the Dynamic Programming Principle and we search for the optimal control policy function, we assume that the controller fix the policy function on $[t, T]$ and follows this policy function on $[t, T]$. Changes in the control policy function are not allowed in the classical setting of dynamic optimization problems. It should be clear that time-inconsistency of the optimal solution has far-reaching consequences since it contradicts the classical notion of optimality and undermines the classical reasoning behind dynamic optimization.

When we consider optimization problems of the form (3.1) the Bellman's Principle of Optimality breaks down and Bellman equation cannot be derived. Let us investigate what happens if we try to repeat the same steps as in the previous section and try to derive the recursive equation for the optimal value function. Let π^* denote the optimal control strategy determined by the optimal control policy function. The optimal control policy function is determined by solving (3.1) for (t, x) and it depends on the initial pair (t, x) . We use the notation $\pi_{t,x}^*$. We define the optimal value function

$$V(t, x) = \mathbb{E}_{t,x} \left[\int_t^T C(t, x, s, X^{\pi_{t,x}^*}(s), \pi_{t,x}^*(s)) ds + G(t, x, X^{\pi_{t,x}^*}(T)) \right],$$

and we still have the relation

$$\begin{aligned} V(t, x) &= \mathbb{E}_{t,x} \left[\int_t^{t+h} C(t, x, s, X^{\pi_{t,x}^*}(s), \pi_{t,x}^*(s)) ds \right. \\ &\quad \left. + \int_{t+h}^T C(t, x, s, X^{\pi_{t,x}^*}(s), \pi_{t,x}^*(s)) ds + G(t, x, X^{\pi_{t,x}^*}(T)) \right]. \end{aligned}$$

However, this time

$$\begin{aligned}
& V(t+h, X^\pi(t+h)) \\
&= \mathbb{E} \left[\int_{t+h}^T C(t+h, X^\pi(t+h), s, X^{\pi_{t+h, X^\pi(t+h)}}^*(s), \pi_{t+h, X^\pi(t+h)}^*(s)) ds \right. \\
&\quad \left. + G(t+h, X^\pi(t+h), X^{\pi_{t+h, X^\pi(t+h)}}^*(T)) | \mathcal{F}_{t+h} \right] \\
&\neq \mathbb{E} \left[\int_{t+h}^T C(t, x, s, X^{\pi_{t,x}^*}(s), \pi_{t,x}^*(s)) ds + G(t, x, X^{\pi_{t,x}^*}(T)) | \mathcal{F}_{t+h} \right].
\end{aligned}$$

We can conclude that the time-consistency of the optimal solution fails when the utilities C and G in the objective (3.1) change as time t goes on or the state process X^π changes its value. The dependence of the intermediate and terminal utility on the initial pair (t, x) , which describes the point in time t and the value of the state process at time t , is not a theoretical sophistication. We now present two examples from finance and insurance, in which we give motivation why we would like to consider time-dependent and state-dependent utility functions in optimization problems.

Problem 1. In Example 1 we consider an investor with constant risk aversion coefficient γ who chooses the investment strategy to maximize the expected exponential utility from the terminal wealth. The optimal control strategy/the optimal control policy function takes the form $\pi^*(t) = \frac{\mu}{\gamma\sigma^2}$.

The assumption that the risk aversion of the investor remains constant over the whole investment period $[0, T]$ might be questioned. There are a lot of arguments for assuming a time-varying risk aversion. In a bull market investors are willing to take more risk, which should be modeled with a lower risk aversion coefficient, whereas in a bear market investors are willing to take less risk, which should be modeled with a higher risk aversion coefficient. Hence, the risk aversion coefficient depending on the state of economy should be used in dynamic portfolio selection problems, see Gordon and St-Amour (2000) and Kwak et al. (2014). There are also strong empirical evidences that the degree of risk aversion depends on prior gains and losses, on wealth in general. After a gain on a prior gamble people are more risk seeking than

usual, while after a prior loss they become more risk averse. The observation that the risk aversion goes down after a prior gain is called the house money effect, and it reflects gamblers increased willingness to bet when ahead, see Thaler and Johnson (1990). Consequently, we should investigate a portfolio selection problem for an investor with risk aversion coefficient depending on the investor's current wealth. We should investigate the optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[- e^{-\gamma(x)X^{\pi}(T)} \right], \quad (3.2)$$

where $\gamma : \mathbb{R} \mapsto (0, \infty)$ is a function of wealth x . Since the optimal control strategy/the optimal control policy function for the investor who maximizes the expected exponential utility of the terminal wealth depends on the risk aversion coefficient, we can easily see that the optimal investment strategy for the investor with risk aversion $\gamma(x)$, who has at time t the available wealth in the amount of x , will no longer be optimal when the wealth process changes its value at some latter point $s > t$ and the investors' risk aversion becomes $\gamma(X^{\pi}(s))$.

We remark that Dong and Sircar (2014) and Delong (2017) study exponential utility maximization problem for an investor with wealth-dependent risk aversion. We also study the optimization problem (3.2) in Example 4 and we present the solution in an explicit form for a specific case. \square

Problem 2. In finance and insurance cash flows are usually discounted with exponential discounting functions. In the classical setting with exponential discounting functions we deal with the optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T e^{-\rho(s-t)} C(X^{\pi}(s), \pi(s)) ds + e^{-\rho(T-t)} G(X^{\pi}(T)) \right],$$

which are equivalent to the optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T e^{-\rho s} C(X^{\pi}(s), \pi(s)) ds + e^{-\rho T} G(X^{\pi}(T)) \right]. \quad (3.3)$$

We can notice that the discounted utilities in (3.3) do not depend on the initial t , hence the optimal solution π^* to (3.3) is time-consistent.

When we investigate optimization problems with exponential discounting functions, we in fact assume that the investor assigns the same discount factor (the same weighting factor) at time t_1 and time $t_2 > t_1$ to value the cash flow at time $t_3 > t_2$. Consequently, the decision rules made at time t remain optimal at latter time $s > t$. However, experimental studies show that our decisions may change as the time passes on. It is well-known that people prefer two oranges in 21 days to one orange in 20 days, but they also prefer one orange now to two oranges tomorrow. Such a feature is called the common difference effect and it cannot be modelled with exponential discounting functions. In the economic literature we can find strong evidences that people discount the future income with non-constant rates of time preferences and the real-life rates of time preference tend to decline in time. In other words, people's valuation tends to decrease rapidly for short period delays and less rapidly for longer period delays, see Loewenstein and Prelec (1992) and Luttmer and Mariotti (2003). Such a feature cannot be described by exponential discounting, but it can be described by hyperbolic discounting.

Let us use a general discounting function ϕ and consider the optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T \phi(s-t) C(s, X^\pi(s), \pi(s)) ds + \phi(T-t) F(X^\pi(T)) \right].$$

The dependence of the discounted utilities on time t cannot be removed and we end up with time-inconsistent optimization problems.

We point out that Ekeland and Lazrak (2006), Ekeland and Pirvu (2008), Ekeland et al. (2012), Alia et al. (2017) and Marin-Solano and Navas (2010) consider investment and consumption problems for agents with general discounting functions. \square

Clearly, time-inconsistency of the optimal solution may arise in many other cases, not just when we deal with time-varying utilities. In the previous section, when deriving (2.9), we point out that we also need the property of conditional expectations to derive the recursive equation for the optimal value

function. Let us consider the family of optimization problems of the form:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T C(s, X^\pi(s), \pi(s)) ds + F(X^\pi(T)) + G(\mathbb{E}_{t,x}[X^\pi(T)]) \right]. \quad (3.4)$$

The term G in the objective (3.4) is a non-linear function of the expected value of the controlled process at the terminal time T . The Dynamic Programming Principle and the time-consistency of the optimal solution fail simply due to the fact that the property of conditional expectation does not hold. Please note that we have the inequality

$$\mathbb{E}_{t,x} \left[G(\mathbb{E}[X^\pi(T)|\mathcal{F}_{t+h}]) \right] \neq G(\mathbb{E}_{t,x}[X^\pi(T)]),$$

which prevents us from repeating the reasoning which lead us to (2.9). We also have optimization problems of the form (3.4) in finance and insurance.

Problem 3. It is well-known that variance is a time-inconsistent risk measure and does not satisfy the property of conditional expectations. Since the work of Henry Markowitz, mean-variance optimization has been one of the key optimization problems considered in finance and insurance. If we want to use a mean-variance risk measure in dynamic portfolio selection, we have to study the optimization problems:

$$\sup_{\pi} \left\{ \mathbb{E}_{t,x} [X^\pi(T)] - \gamma \left(\mathbb{E}_{t,x} [|X^\pi(T)|^2] - \left(\mathbb{E}_{t,x} [X^\pi(T)] \right)^2 \right) \right\}, \quad (3.5)$$

where the last term is a non-linear function of the expected wealth. The objective (3.5) fits (3.4). As noticed by Björk et al. (2014), in real-life applications we should consider the risk aversion coefficient γ which depends on wealth x . A wealth-dependent risk aversion $\gamma(x)$ would introduce the second source of time-inconsistency to our optimization problem, which we discuss in Problem 1.

Let us remark that Hu et al. (2012), Zeng and Li (2011), Björk et al. (2014) and Kronborg and Steffensen (2015) study mean-variance optimization problems. \square .

4 Equilibrium strategies and extended Hamilton-Jacobi-Bellman equation

We now focus on time-inconsistent optimization problems of the form (3.1), i.e. in this section we study the optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[\int_t^T C(t, x, s, X^\pi(s), \pi(s)) ds + G(t, x, X^\pi(T)) \right]. \quad (4.1)$$

We know that the Bellman's Principle of Optimality does not hold for (4.1) and we cannot use the arguments from Section 2 to define the optimal solution in the classical sense. In this section we explain how to define a solution to the time-inconsistent dynamic optimization problem (4.1).

We have two simple approaches to handle the time-inconsistency. Let $\pi_{t,x}^*$ denote the optimal control policy function found by solving (4.1) for the initial pair (t, x) . The policy $\pi_{t,x}^*$ is found by solving the HJB equation (2.11) with the utilities C and G valued and fixed at (t, x) , i.e. $\pi_{t,x}^*$ is found by solving the HJB equation:

$$\begin{aligned} \sup_{\pi} \{ \mathcal{L}^\pi V^{t,x}(s, y) + C(t, x, s, y, \pi) \} &= 0, \quad (s, y) \in [t, T) \times \mathbb{R}, \\ V^{t,x}(T, y) &= G(t, x, y), \quad y \in \mathbb{R}. \end{aligned} \quad (4.2)$$

We next define

$$\pi_{t,x}^*(s, y) = \arg \max_{\pi} \{ \mathcal{L}^\pi V^{t,x}(s, y) + C(t, x, s, y, \pi) \}, \quad (s, y) \in [t, T) \times \mathbb{R}.$$

We can use two types of solutions:

- Pre-commitment solution: Use $\pi_{0,x_0}^*(t, x)$ at time t given that $X(t) = x$, where x_0 is the initial value of the state process X at time 0,
- Naive solution: Use $\pi_{t,x}^*(t, x)$ at time t given that $X(t) = x$.

Example 3. We consider Problem 1 and the exponential utility maximization problem for an investor with wealth-dependent risk aversion. We

want to solve the optimization problems:

$$\sup_{\pi} \mathbb{E}_{t,x} \left[- e^{-\gamma(x)X^{\pi}(T)} \right].$$

Using the results from Example 2 we can conclude that the pre-commitment solution is $\pi^*(t) = \frac{\mu}{\gamma(x_0)\sigma^2}$, where x_0 is the initial wealth of the investor at time 0, and the naive solution is $\pi^*(t) = \frac{\mu}{\gamma(X^{\pi^*(t)})\sigma^2}$. \square

The advantage of the pre-commitment and naive solution is that they are derived by solving classical HJB equations and are based on the notion of optimality described in Section 2. The disadvantage of the pre-commitment and naive solution is that they ignore the key feature of the dynamic optimization problem (4.1) which is the time-varying utilities C and G . When we use the pre-commitment solution we assume that the controller who solves the dynamic optimization problem (4.1) at time $t = 0$ can force the future controllers to use his/her strategy, even though this strategy will not be the optimal strategy for the future controllers with different utilities. The naive solution tries to glue the strategies which are optimal for all controllers. However, the naive solution to the optimization problem (4.1) for the initial pair (t, x) is derived under the assumption that all future controllers will use the same utilities $C(t, x, \dots)$ and $G(t, x, \dots)$ as the controller with wealth x at time t who search for the optimal decision rule at time t . Hence, future changes in the utilities are still not modelled. Clearly, we would like to find a solution to the dynamic optimization problem (4.1) which takes into account that the utilities C and G are time-varying and investor's preferences are changing. In other words, we would like to find the optimal decision rule for the controller with utilities $C(t, x, \dots)$ and $G(t, x, \dots)$ given the knowledge that the future controllers may have different utilities, depending on the future wealth and time, and may apply different decision rules in accordance with their utilities. Such a solution is called a sophisticated solution and it requires a different concept of optimality.

We take a game-theoretic approach. Let us consider a game played by a continuum of agents during the time interval $[0, T]$. The agent at time

t only chooses the strategy at time t . When the agents' utilities are constant, the future agents will solve the remaining part of the optimization problem faced by the agent at time t . However, when the agents' utilities are time-varying, the future agents will not solve the remaining part of the optimization problem faced by the agent at time t since the objective function changes constantly. Indeed, the agent who has the wealth in the amount of x at time t aims at maximizing the objective:

$$\mathbb{E}_{t,x} \left[\int_t^T C(t, x, s, X^\pi(s), \pi(s)) ds + G(t, x, X^\pi(T)) \right]. \quad (4.3)$$

The objective changes since the utilities C and G change with time and available wealth. Consequently, the agents face different optimization problems. In this framework of a game played by a continuum of agents, the reward to the agent at time t , i.e. the value of the objective function (4.3), depends on the strategy chosen by himself/herself and the strategies chosen by all future agents. Hence, the agent at time t plays with the other agent who will make decisions after time t . The question is what should the agent at time t do, taking into account the decision-making of the future agents? If the agent at time t follows the naive approach and chooses the best strategy according to his/her preferences (i.e. he/she solves the classical optimization problem with the fixed utilities $C(t, x, \dots)$ and $G(t, x, \dots)$ by applying the Bellman's Principle of Optimality and the Dynamic Programming Principle), then his/her optimal strategy will not be adopted by the future agents who will have different utilities, will solve different optimization problems and will apply different optimal strategies. Consequently, the true reward to the agent at time t will be lower than the reward resulting from his/her naive optimization process. It seems reasonable to assume that the agent at time t should sacrifice the short-term benefit to gain in the long-term. We look for a strategy in the sub-game perfect Nash equilibrium.

In game theory, the Nash equilibrium is a solution concept of a non-cooperative game involving two or more players in which each player knows the equilibrium strategies of the other players, and the player does not benefit

by changing only his/her own strategy. If each player has chosen a strategy and no player can improve his/her reward by changing the strategy while the other players keep their strategies unchanged, then the current set of strategies constitutes the Nash equilibrium. A sub-game perfect Nash equilibrium is an improved version of the Nash equilibrium which eliminates non-rational decisions in sequential games such as we have here in the time period $[0, T]$.

As before, let V^π denote the objective function/the reward under a control strategy π , i.e.

$$V^\pi(t, x) = \mathbb{E}_{t,x} \left[\int_t^T C(t, x, s, X^\pi(s), \pi(s)) ds + G(t, x, X^\pi(T)) \right].$$

In a discrete time model, where the strategies are chosen at discrete times $0, h, \dots, t, t+h, \dots, T-h$ and kept fixed in between, the equilibrium strategy is well understood. It is known that the strategy in a Nash equilibrium is the best response to all other strategies in that equilibrium. Hence, the equilibrium strategy can be derived with the procedure:

- Let the agent at time $T-h$ optimize the objective functional $V^\pi(T-h, x)$ over π_{T-h} for all x ,
- Let the agent at time $T-2h$ optimize the objective functional $V^\pi(T-2, x)$ over π_{T-2} for all x , given the knowledge that the agent at time $T-h$ will use π_{T-1}^* ,
- Proceed recursively by induction.

Let us remark that this sequential procedure is similar to the procedure described in Section 2, when we derive a recursive equation for the value function (2.9) and a step-by-step procedure for finding the optimal strategies. However, for a time-inconsistent optimization problem we do not have a simple relationship between the value function in one period and the value function in the next period. Let us refer the reader to Yong (2012) where a multi-person game in a discrete time is studied and a step-by-step procedure for finding the optimal strategies for a time-inconsistent problem is established.

We can formalize the definition of an equilibrium strategy in a discrete-time model. Let us recall that a strategy is a Nash equilibrium strategy if no agent can do better by unilaterally changing his/her strategy (knowing the strategies of the other agents)

Definition 4.1. *Let us consider a strategy π^* . Choose an arbitrary point $(t, x) \in \{0, h, \dots, T - h\} \times \mathbb{R}$ and any strategy π . We define a new strategy*

$$\pi_\delta(s, y) = \begin{cases} \pi(y), & s = t, \\ \pi^*(s, y), & s = t + h, \dots, T - h, \quad y \in \mathbb{R} \end{cases}$$

If

$$\sup_{\pi} V^{\pi_\delta}(t, x) = V^{\pi^*}(t, x)$$

then π^* is called an equilibrium strategy and $V^{\pi^*}(t, x)$ is called the equilibrium value function corresponding to the equilibrium strategy π^* .

Unfortunately, we cannot apply the above definition in continuous-time models. The reason is that in a continuous-time model a change in the control strategy at time t does not affect the controlled process (2.1) and the objective function (4.3). The definition of an equilibrium strategy in a continuous-time model is a bit more theoretical, but the main idea behind the equilibrium strategy remains.

Definition 4.2. *Let us consider a strategy π^* . Choose an arbitrary point $(t, x) \in [0, T) \times \mathbb{R}$ and any strategy π . We define a new strategy*

$$\pi_\delta(s, y) = \begin{cases} \pi(s, y), & t \leq s \leq t + \delta, \\ \pi^*(s, y), & t + \delta < s \leq T, \quad y \in \mathbb{R}. \end{cases}$$

If

$$\liminf_{\delta \rightarrow 0} \frac{V^{\pi^*}(t, x, p) - V^{\pi_\delta}(t, x, p)}{\delta} \geq 0,$$

then π^* is called an equilibrium strategy and $V^{\pi^*}(t, x)$ is called the equilibrium value function corresponding to the equilibrium strategy π^* .

We would like to remark that we define an equilibrium strategy in the class of closed-loops control strategies. It is also possible to define an equilibrium strategy in the class of open-loops control strategies, see Hu et al. (2012). In general, the equilibrium strategy in the class of closed-loops control strategies is different from the equilibrium strategy in the class of open-loops control strategies, see Hu et al. (2012) and Alia et al. (2017).

At the beginning of this section we introduce pre-commitment and naive solutions. The solution in a Nash equilibrium is called a sophisticated solution. Let us point out that all three types of solutions, pre-commitment, naive and sophisticated, are different in general. For a comparison of these three types of solutions we refer to Marin-Solano and Navas (2010) who investigate optimal consumption and investment problems.

Interestingly, we can still establish a recursive equation for the objective function under the equilibrium strategy. Consequently, we can derive a version of Hamilton-Jacobi-Bellman equation which characterizes the equilibrium strategy and the equilibrium value function. As expected, the HJB equation for the equilibrium value function for a time-inconsistent optimization problem is much complicated than the HJB equation for the optimal value function for a time-consistent optimization problem. First, let us present the idea behind the HJB equation for the equilibrium value function. For simplicity, we consider the optimization problem (4.1) without the intermediate utility C and with the terminal utility G depending only on x . Our goal is to solve

$$\sup_{\pi} \mathbb{E}_{t,x} \left[G(x, X^{\pi}(T)) \right]. \quad (4.4)$$

Let π^* denote an equilibrium strategy and V denote the equilibrium value function, i.e. $V(t, x) = V^{\pi^*}(t, x)$. We can derive the following recursion:

$$\begin{aligned} V(t, x) &= \mathbb{E}_{t,x} \left[G(x, X^{\pi^*}(T)) \right] \\ &= \mathbb{E}_{t,x} \left[V(t+h, X^{\pi^*}(t+h)) \right. \\ &\quad \left. - \left(W(t, X^{\pi^*}(t+h), X^{\pi^*}(t+h)) - W(t, X^{\pi^*}(t+h), x) \right) \right], \quad (4.5) \end{aligned}$$

where we introduce the auxiliary value function:

$$W(t, x, y) = \mathbb{E}_{t,x} \left[G(y, X^{\pi^*}(T)) \right]. \quad (4.6)$$

The function W gives the objective function for the optimization problem (4.4) under the equilibrium strategy with the utility G depending on an auxiliary parameter y . When we compare the recursion (2.9) for the optimal value function with the recursion (4.5) for the equilibrium value function, we can see that we now have one additional term $W(t, X^{\pi^*}(t+h), X^{\pi^*}(t+h)) - W(t, X^{\pi^*}(t+h), x)$ which describes the change in the equilibrium value function resulting from changes in the preferences.

If we divide the equation (4.5) by h and let $h \rightarrow 0$, we end up with a so-called extended Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} \sup_{\pi} \left\{ \mathcal{L}^{\pi} V(t, x) - (\mathcal{M}^{\pi} W(t, x, x) - \mathcal{L}^{\pi} W(t, x, x)) \right\} &= 0, \quad (t, x) \in [0, T) \times \mathbb{R}, \\ V(T, x) &= G(x, x), \quad x \in \mathbb{R}, \end{aligned} \quad (4.7)$$

$$\begin{aligned} \mathcal{L}^{\pi^*} W(t, x, y) &= 0, \quad (t, x) \in [0, T) \times \mathbb{R}, \quad y \in \mathbb{R}, \\ W(T, x, y) &= G(y, x), \quad x \in \mathbb{R}, \quad y \in \mathbb{R}, \end{aligned} \quad (4.8)$$

where

$$\begin{aligned} \mathcal{M}^{\pi} f(t, x, y) &= \mathcal{L}^{\pi} f(t, x, y) \\ &\quad + \mu(t, x, \pi) f_y(t, x, y) + \frac{1}{2} \sigma^2(t, x, \pi) f_{yy}(t, x, y) \\ &\quad + \sigma^2(t, x, \pi) f_{xy}(t, x, y). \end{aligned}$$

The term $\mathcal{M}^{\pi} W(t, x, x)$ in (4.7) should be understood as $\mathcal{M}^{\pi} W(t, x, y)|_{y=x}$, and the generator \mathcal{L}^{π} is applied on W by treating the last variable as fixed, see (2.2) for the definition of \mathcal{L}^{π} . The equation (4.8) follows from Feynman-Kac formula applied to the auxiliary value function (4.6).

We now consider the general optimization problem (4.1). We present the verification theorem and the extended HJB equation, see Definition 4.4 in Björk and Murgoci (2014).

Theorem 4.1. *Let the operator \mathcal{L} be defined in (2.2) and let the operator \mathcal{M} be defined as*

$$\begin{aligned}\mathcal{M}^\pi f(t, x, r, y) &= \mathcal{L}^\pi f(t, x, r, y) \\ &+ f_r(t, x, r, y) + \mu(t, x, \pi) f_y(t, x, r, y) + \frac{1}{2} \sigma^2(t, x, \pi) f_{yy}(t, x, r, y) \\ &+ \sigma^2(t, x, \pi) f_{xy}(t, x, r, y).\end{aligned}$$

The operators \mathcal{L}^π and \mathcal{M}^π act on $f \in \mathcal{C}^{1,2,1,2}([0, T] \times \mathbb{R} \times [0, T] \times \mathbb{R})$. Assume there exist functions $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R})$, $W \in \mathcal{C}^{1,2,1,2}([0, T] \times \mathbb{R} \times [0, T] \times \mathbb{R})$, $U \in \mathcal{C}^{1,2,1,2,0}([0, T] \times \mathbb{R} \times [0, T] \times \mathbb{R} \times [0, T])$ and a strategy π^ which solve the system of HJB equations:*

$$\begin{aligned}\sup_{\pi} \left\{ \mathcal{L}^\pi V(t, x) + C(t, x, t, x, \pi) - (\mathcal{M}^\pi W(t, x, t, x) - \mathcal{L}^\pi W(t, x, t, x)) \right. \\ \left. - \int_t^T (\mathcal{M}^\pi U(t, x, t, x, s) - \mathcal{L}^\pi U(t, x, t, x, s)) ds \right\} = 0, \quad (t, x) \in [0, T] \times \mathbb{R}, \\ V(T, x) = G(T, x, x), \quad x \in \mathbb{R},\end{aligned}\tag{4.9}$$

$$\begin{aligned}\mathcal{L}^{\pi^*} W(t, x, r, y) &= 0, \quad (t, x) \in [0, T] \times \mathbb{R}, \\ W(T, x, r, y) &= G(r, y, x), \quad x \in \mathbb{R},\end{aligned}\tag{4.10}$$

$$\begin{aligned}\mathcal{L}^{\pi^*} U(t, x, r, y, s) &= 0, \quad (t, x) \in [0, s] \times \mathbb{R}, \\ U(s, x, r, y, s) &= C(r, y, s, x, \pi^*(s, x)), \quad x \in \mathbb{R},\end{aligned}\tag{4.11}$$

for all $(r, y) \in [0, T] \times \mathbb{R}$ and $s \in [0, T]$. The strategy π^ is an equilibrium strategy for the optimization problem (4.1) and $V(t, x) = V^{\pi^*}(t, x)$ is the equilibrium value function corresponding to the equilibrium strategy π^* . Moreover, $V(t, x) = W(t, x, t, x) + \int_t^T U(t, x, t, x, s) ds$.*

Let us remark that the operators in (4.9)-(4.11) should be understood as in (4.7). From Theorem 4.1 we can deduce probabilistic representations of

the unknown functions. By Feynman-Kac formula we have:

$$\begin{aligned}
V(t, x) &= \mathbb{E}_{t,x} \left[\int_t^T C(t, x, s, X^{\pi^*}(s), \pi^*(s, X^{\pi^*}(s))) ds + G(t, x, X^{\pi^*}(T)) \right], \\
W(t, x, r, y) &= \mathbb{E}_{t,x} \left[G(r, y, X^{\pi^*}(T)) \right], \\
U(t, x, r, y, s) &= \mathbb{E}_{t,x} \left[C(r, y, s, X^{\pi^*}(s), \pi^*(s, X^{\pi^*}(s))) ds \right].
\end{aligned}$$

The extended HJB equation (4.9)-(4.11) is a system of three equations. The equilibrium strategy is derived from the first equation (4.9) which can be solved if the functions W and U are known. The functions W and U are characterized with the equations (4.10)-(4.11) which can be solved if the equilibrium strategy is found. We can look at the system of equations (4.9)-(4.11) as if it was a fixed-point equation for the equilibrium strategy. We can solve the system in the following way:

- Choose an arbitrary strategy $\pi^{*,1}$,
- Solve the equations (4.10)-(4.11) and find W and U ,
- Solve the equation (4.9) with the functions W and U from the previous step and find a new strategy $\pi^{*,2}$
- Iterate the procedure until convergence for the sequence $(\pi^{*,k})_{k=1,2,\dots}$ is reached.

Example 4. We consider Problem 1. We deal with the optimization problem (4.4) with $G(y, x) = -e^{-\gamma(y)x}$. From (4.7)-(4.8) and Theorem 4.1 we can conclude that the equilibrium strategy and the equilibrium value

function are characterized with the HJB equations:

$$\begin{aligned} \sup_{\pi} \left\{ V_t(t, x) + \pi \mu V_x(t, x) + \frac{1}{2} \pi^2 \sigma^2 V_{xx}(t, x) - \pi \mu W_y(t, x, x) \right. \\ \left. - \frac{1}{2} \pi^2 \sigma^2 W_{yy}(t, x, x) - \pi^2 \sigma^2 W_{xy}(t, x, x) \right\} = 0, \quad (t, x) \in [0, T) \times \mathbb{R}, \\ V(T, x) = -e^{-\gamma(x)x}, \quad x \in \mathbb{R}, \end{aligned} \quad (4.12)$$

$$\begin{aligned} W_t(t, x, y) + \pi^*(t, x) \mu W_x(t, x, y) \\ + \frac{1}{2} (\pi^*(t, x))^2 \sigma^2 W_{xx}(t, x, y) = 0, \quad (t, x) \in [0, T) \times \mathbb{R}, y \in \mathbb{R}, \\ W(T, x, y) = -e^{-\gamma(y)x}, \quad x \in \mathbb{R}, y \in \mathbb{R}. \end{aligned} \quad (4.13)$$

Let us assume that $T = 1$, $\mu = 0.5$, $\sigma = 0.1$ and $\gamma(x) = 0.3 + 0.2\Phi(-(x-115))$, where Φ denotes the standard normal distribution function. The risk aversion as a function of wealth is presented in Figure 1. The higher the wealth, the lower the coefficient of risk aversion. We solve the HJB equations (4.12)-(4.13) by using the fixed point procedure and the implicit difference scheme. The equilibrium strategy and the naive strategy are presented in Figure 2. Let us recall that the naive strategy is given by $\pi(t, x) = \frac{\mu}{\sigma^2 \gamma(x)}$, see Example 3.

The equilibrium investment strategy is similar in shape (as a function of wealth) to the naive investment strategy but the equilibrium investment strategy does not coincide with the naive investment strategy, see Figure 2. As expected, for both the equilibrium strategy and the naive strategy: the higher the wealth, the higher the amount of money invested in the risky stock (since the risk aversion decreases as the wealth increases). However, the equilibrium investment strategy increases with wealth slower than the naive investment strategy. The amount of money invested in the risky stock given by the equilibrium strategy is lower than the amount of money given by the naive strategy, especially for initial times t , and this discrepancy decreases as time t approaches maturity T , see Figure 2. This observation agrees with intuition. If the available wealth is high, then the naive solution tells us to invest a high amount of money in the risky stock since the risk aversion is low. However, the naive solution of the optimization problem assumes

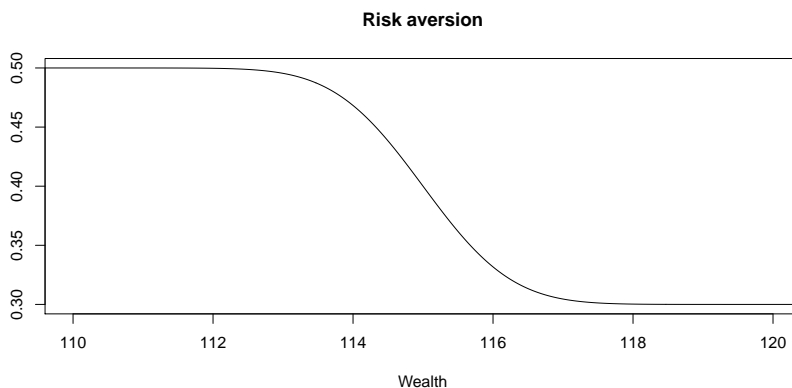


Figure 1: The coefficient of risk aversion as a function of wealth.

that all future investors will have low coefficients of risk aversion or that the investor at time t can commit all future investors to apply his/her strategy. The naive solution does not take into account that the wealth may decrease in the future, the coefficient of risk aversion may increase and the future investors may prefer to invest lower amounts of money in the risky stock. Consequently, the strategy chosen by the naive agent at time t will not be adopted by the future agents. The equilibrium strategy at time t does take into account investment decisions preferred by the future investors who may have different risk preferences and may opt for lower allocations in the risky stock. The sophisticated solution of the optimization problem tells us to invest less money in the risky stock compared to the naive solution. As time t approaches maturity T , the probability that the wealth decreases before maturity and the future investors will switch to lower allocations in the risk stock becomes lower. Hence, the investor close to maturity, who follows the sophisticated solution, can invest higher amounts of money in the risk stock and his/her investment strategy becomes closer to the naive strategy. \square

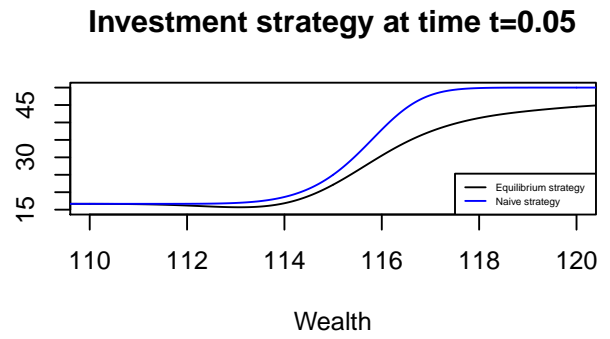
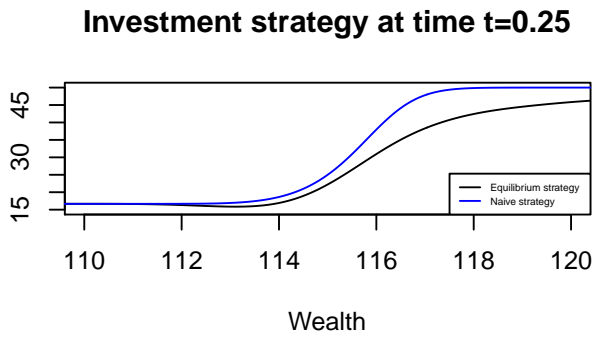
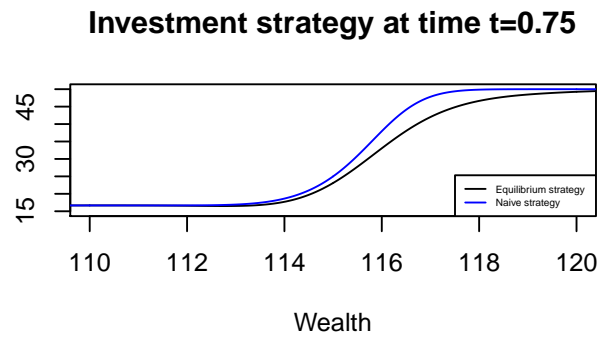
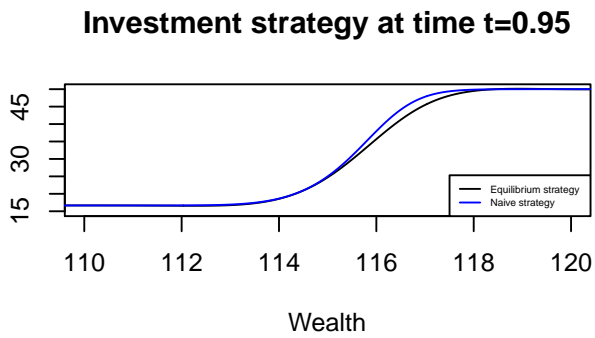


Figure 2: The equilibrium strategy and the naive strategy (the amounts of money invested in the risky stock).

5 Conclusions

In this paper we have studied time-inconsistent stochastic optimal control problems. We have discussed the concepts of time-consistency, time-inconsistency, optimal strategy, Nash equilibrium strategy and extended Hamilton-Jacobi-Bellman equation. We have given three examples of time-inconsistent dynamic optimization problems which can arise in insurance and finance and we have presented the solution for exponential utility maximization problem with wealth-dependent risk aversion.

References

- Alia I, Chighoub F, Khelfallah N, Vives J (2017) Time-consistent investment and consumption strategies under a general discount function. Preprint
- Björk T, Murgoci A (2014) A theory of Markovian time-inconsistent stochastic control in discrete time. *Finance and Stochastics* 18: 545-592
- Björk T, Khapko M, Murgoci A (2017) On time-inconsistent stochastic control in continuous time. *Finance and Stochastics* 21:331-360
- Björk T, Murgoci A, Zhou X (2014) Mean-variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance* 24:1-24
- Carmona R (2009) *Indifference Pricing: Theory and Applications*. Princeton University Press
- Delong L (2017) Optimal investment for insurance company with exponential utility and wealth-dependent risk aversion coefficient. Preprint
- Delong L, Chen A (2016) Asset allocation, sustainable withdrawal, longevity risk and non-exponential discounting. *Insurance: Mathematics and Economics* 71:342-352

- Dong Y, Sircar R (2014) Time-inconsistent portfolio investment problems. *Stochastic Analysis and Applications* 100:239-281
- Ekeland I, Lazrak A (2006) Being serious about non-commitment: subgame perfect equilibrium in continuous time. Preprint
- Ekeland I, Pirvu T (2008) Investment and consumption without commitment. *Mathematical Financial Economics* 2:57-86
- Ekeland I, Mbodji O, Pirvu T (2012) Time-consistent portfolio management. *SIAM Journal of Financial Mathematics* 3:1-32
- Fleming W, Rishel R (1975) *Deterministic and Stochastic Optimal Control*. Springer
- Gordon S, St-Amour P (2000) A preference regime model of bull and bear markets. *American Economic Review* 90:1019-1033
- Hu Y, Jin H, Zhou XY (2012) Time-inconsistent stochastic linear-quadratic control. *SIAM Journal on Control and Optimization* 50:1548-1572
- Kronborg M, Steffensen M (2015) Inconsistent investment and consumption problems. *Applied Mathematics and Optimization* 71:473-515
- Kwak M, Pirvu T, Zhang H (2014) A multiperiod equilibrium pricing model. *Journal of Applied Mathematics* 2014:1-14
- Loewenstein G, Prelec D (1992) Anomalies in intertemporal choices: evidence and an interpretation. *The Quarterly Journal of Economics* 107:573-597
- Luttmer EGJ, Mariotti T (2003) Subjective discounting in an exchange economy. *Journal of Political Economy* 111:959-989
- Marin-Solano J, Navas J (2010) Consumption and portfolio rules for time-inconsistent investors. *European Journal of Operational Research* 201:860-872

- Øksendal B, Sulem A (2004) Applied Stochastic Control of Jump Diffusions. Springer
- Pham H (2009) Continuous-time Stochastic Control and Optimization with Financial Applications. Springer
- Thaler R, Johnson E (1990) Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Management Science* 36:643-660.
- Zeng Y, Li Z (2011) Optimal time-consistent investment and reinsurance policies for mean-variance insurers. *Insurance: Mathematics and Economics* 49:145-154.
- Yong J (2012) Time-inconsistent optimal control problems and the equilibrium HJB equation. *American Institute of Mathematical Sciences* 2: 271-329
- Yong J, Zhou XY (1999) Hamiltonian Systems and HJB Equations. Springer